PLPSTA31 – Statistiques 2
Tests d'hypothèses statistiques

Estimation ponctuelle – Distribution empirique des moyennes

Mode d'emploi des questionnaires à choix multiples

Pour démarrer un exercice, il faut cliquer sur Début.
Pour afficher le score obtenu au questionnaire et les réponses correctes, il faut cliquer sur Fin, puis sur le bouton Réponse :

- le signe ✔ indique que la réponse donnée est correcte ;
- le signe ✘ indique que la réponse donnée est incorrecte ; les réponses correctes manquantes sont marquées par ●.

Pour certaines questions, les réponses sont détaillées : cliquer sur un des symboles verts ✔ ou ● pour accéder aux compléments de réponse.

Attention : il peut y avoir plusieurs propositions correctes dans une liste. Il faut toutes les cocher pour que la réponse à la question soit considérée comme exacte.

Exercice 1

Chez les employés d'une grande entreprise, on veut estimer la proportion de salariés femmes, le salaire moyen des employés et l'écart-type des salaires.

Sur un échantillon de 32 employés tirés au sort dans la population, on observe 48% de femmes, un salaire moyen égal à 1200 (euros) et un écart-type « corrigé » des salaires égal à 340 (euros).

Début Parmi les propositions suivantes, cocher celles qui sont correctes.

1. La population étudiée est :

- [ ] les employés femmes de l'entreprise
- [ ] 32 employés tirés au sort
- [ ] les employés de l'entreprise

2. Les variables étudiées dans la population sont :

- [ ] la proportion de femmes
- [ ] le sexe
- [ ] le salaire moyen des employés
- [ ] le salaire des employés
- [ ] le salaire des employés femmes

3. La variable sexe est une variable

- [ ] qualitative
- [ ] quantitative discrète
- [ ] quantitative continue

4. La proportion d'employés femmes de l'entreprise est :

- [ ] notée p
- [ ] notée f
- [ ] inconnue
- [ ] égale à 48%

5. Le salaire moyen des employés de l'entreprise est :

- [ ] noté μ
- [ ] noté $\bar{x}$

6. Le salaire moyen des 32 employés de l'échantillon qui a été tiré au sort est :

- [ ] noté μ
- [ ] noté $\bar{x}$

7. Le salaire moyen des employés de l'entreprise :

- [ ] est égal à 1200 euros
- [ ] est estimé à 1200 euros
- [ ] sa vraie valeur reste inconnue

8. la moyenne 1200 est la valeur observée (sur l'échantillon tiré au sort) de la statistique :

- [ ] moyenne empirique $\bar{X}_n$
- [ ] fréquence empirique F_n
- [ ] écart-type empirique S_n

Fin Score : Réponse Cliquer sur ✔ ou ● pour accéder aux compléments de réponse.

Exercice 2

Début

On étudie la variable *note à un examen* X, variable quantitative de moyenne μ et d'écart-type σ, définie dans une population donnée d'étudiants.

1. Déterminer la forme de la distribution de la moyenne empirique $\bar{X}_n$ dans les situations suivantes :

(a) $n = 16$, σ inconnu, loi de X normale dans la population :

☐ loi normale ☐ loi approximativement normale ☐ loi exacte inconnue

(b) $n = 12$, $\sigma = 3$, loi de X inconnue dans la population :

☐ loi normale ☐ loi approximativement normale ☐ loi exacte inconnue

(c) $n = 31$, σ inconnu, loi de X normale dans la population :

☐ loi normale ☐ loi approximativement normale ☐ loi exacte inconnue

(d) $n = 31$, σ inconnu, loi de X inconnue dans la population :

☐ loi normale ☐ loi approximativement normale ☐ loi exacte inconnue

2. Quelle est la relation entre la moyenne de la distribution de $\bar{X}_n$ et la moyenne μ de la population :

☐ aucune relation ☐ la moyenne de $\bar{X}_n$ est égale à μ

☐ la moyenne μ est le centre de la distribution de $\bar{X}_n$ ☐ la moyenne de $\bar{X}_n$ est inconnue pour $n < 30$

3. La dispersion des moyennes d'échantillons autour de leur moyenne est mesurée par l'écart-type de la distribution de $\bar{X}_n$. Cette dispersion :

☐ est d'autant plus élevée que l'écart-type σ de la population est grand

☐ est d'autant plus faible que l'écart-type σ de la population est grand

☐ est d'autant plus faible que n est grand

Fin | Score : | Réponse | Cliquer sur ✔ ou ● pour accéder aux compléments de réponse.

Retour

Les réponses aux questionnaires

Réponse :

La population est composée de tous les employés (hommes, femmes) de l'entreprise. Les femmes constituent une catégorie particulière d'employés dans la population.

Les 32 employés ont été tirés au sort dans la population. Ils constituent l'échantillon ; les données (sexe et salaire) observées sur l'échantillon apportent une information partielle sur la population.

Retour au questionnaire.

Réponse : Il y a deux variables étudiées dans la population des employés : le sexe et le salaire.

La variable Sexe est qualitative à deux modalités homme, femme ; la variable Salaire est quantitative, traitée en général comme une variable continue.

La proportion de femmes n'est pas une variable mais le paramètre (à estimer) de la variable Sexe.

Le salaire moyen est la moyenne μ de la variable Salaire. Ce paramètre est inconnu et à estimer. L'autre paramètre étudié est l'écart-type σ de la variable Salaire. Retour au questionnaire.

Réponse :

On note ici f la fréquence de femmes observée dans l'échantillon, soit 48%.

La proportion p d'employés femmes de l'entreprise est inconnue. On peut l'estimer à 48%, estimation ponctuelle fournie par l'échantillon.

Retour au questionnaire.

Retour

Réponse :

On note $\bar{x}$ le salaire moyen observé sur l'échantillon tiré au sort dans la population. C'est une valeur numérique (1200) qui résume les 32 salaires de l"échantillon.

La lettre μ désigne la moyenne de la population, ici le salaire moyen de l'ensemble des employés de l'entreprise.

Retour au questionnaire.

Retour

Réponse :

- Pour l'étude des salaires de la population, la statistique utilisée pour estimer le paramètre μ est la moyenne empirique $\bar{X}_n$.
  La moyenne observée 1200 est la valeur observée de $\bar{X}_n$ sur l'échantillon ; c'est la valeur estimée de μ. La vraie valeur de μ reste inconnue.
- Pour estimer l'écart-type σ des salaires de l'entreprise, la statistique utilisée est l'écart-type empirique sans biais S_n^*.
  L'écart-type observé 340 est la valeur observée de S_n^*. C'est l'estimation ponctuelle « sans biais » de σ.
  L'écart-type empirique biaisé S_n a pour valeur l'écart-type s de l'échantillon. Cette valeur observée n'est pas donnée dans l'énoncé.
- Pour l'estimation de la proportion p de femmes chez les employés de l'entreprise, la statistique utile est la fréquence empirique de la catégorie femme, que l'on note F_n.
  Sa valeur observée sur l'échantillon est 48% ; la proportion p est estimée à 48%.

Retour au questionnaire.

Retour

Réponse :

L'information sur la valeur de σ (inconnue ou donnée) est sans intérêt pour la question 1.

Pour les paramètres de la loi de la moyenne empirique $\overline{X}_n$:

La moyenne de $\overline{X}_n$ est μ (μ est la moyenne de la variable X) et son écart-type est $\frac{\sigma}{\sqrt{n}}$ (σ est l'écart-type de X).

Pour établir la forme de la distribution de $\overline{X}_n$, on dispose de deux résultats :

– si la loi de X est normale (de moyenne μ et d'écart-type σ), alors la moyenne empirique $\overline{X}_n$ a pour distribution la loi normale (de moyenne μ et d'écart-type $\frac{\sigma}{\sqrt{n}}$).

– Si la loi de X est quelconque (inconnue), la loi de $\overline{X}_n$ est approximativement normale pour n assez grand ($n \geq 30$). Quand la loi de X est inconnue, la loi exacte de $\overline{X}_n$ est inconnue.

Retour au questionnaire.

Retour

Réponse : L'écart-type de $\bar{X}_n$ mesure la dispersion des moyennes de tous les échantillons possibles de taille n (fixée) autour de μ. Il est égal à $\frac{\sigma}{\sqrt{n}}$.

La formule montre que l'écart-type de $\bar{X}_n$ diminue quand n augmente et qu'il augmente avec σ.

Cette formule corrobore les raisonnements que l'on peut faire intuitivement :

Plus les notes sont dispersées dans la population et plus les moyennes de n notes prises au hasard dans la population sont elles aussi dispersées.

Plus n augmente et plus les moyennes des échantillons (moyennes calculées sur n étudiants pris dans la population) se rapprochent de la moyenne μ de la population (moyenne de la totalité des étudiants).

Retour au questionnaire.